

In [mathematics](#), a **matrix** (pl.: **matrices**) is a [rectangular](#) array of [numbers](#) or other [mathematical objects](#) with elements or entries arranged in rows and columns, usually satisfying certain properties of [addition](#) and [multiplication](#).

For example,

$\begin{bmatrix} 1 & 9 & -13 \\ 20 & 5 & -6 \end{bmatrix}$ denotes a matrix with two rows and three columns. This is often referred to as a "two-by-three matrix", a 2×3 matrix, or a matrix of dimension 2×3 .

In [linear algebra](#), matrices are used as [linear maps](#). In [geometry](#), matrices are used for [geometric transformations](#) (for example [rotations](#)) and [coordinate changes](#). In [numerical analysis](#), many computational problems are solved by reducing them to a matrix computation, and this often involves computing with matrices of huge dimensions. Matrices are used in most areas of mathematics and scientific fields, either directly, or through their use in geometry and numerical analysis.

[Square matrices](#), matrices with the same number of rows and columns, play a major role in matrix theory. The [determinant](#) of a square matrix is a number associated with the matrix, which is fundamental for the study of a square matrix; for example, a square matrix is [invertible](#) if and only if it has a nonzero determinant and the [eigenvalues](#) of a square matrix are the roots of its [characteristic polynomial](#), $\det(\lambda I - A)$.

Matrix theory is the [branch of mathematics](#) that focuses on the study of matrices. It was initially a sub-branch of linear algebra, but soon grew to include subjects related to [graph theory](#), [algebra](#), [combinatorics](#) and [statistics](#).

Definition

A matrix is a rectangular array of [numbers](#) (or other mathematical objects), called the "entries" of the matrix. Matrices are subject to standard [operations](#) such as [addition](#) and [multiplication](#).^[1] Most commonly, a matrix over a [field](#) F is a rectangular array of [elements](#) of F .^{[2][3]} A **real matrix** and a **complex matrix** are matrices whose entries are respectively [real numbers](#) or [complex numbers](#). More general types of entries are discussed [below](#). For instance, this is a real matrix:

$$A = \begin{bmatrix} -1.3 & 0.6 \\ 20.4 & 5.5 \\ 9.7 & -6.2 \end{bmatrix}.$$

The numbers (or other objects) in the matrix are called its *entries* or its *elements*. The horizontal and vertical lines of entries in a matrix are respectively called *rows* and *columns*.^[4]

Size

The size of a matrix is defined by the number of rows and columns it contains. There is no limit to the number of rows and columns that a matrix (in the usual sense) can have as long as they are positive integers. A matrix with m rows and n columns is called an $m \times n$ matrix,^[4] or m -by- n matrix,^[5] where m and n are called its *dimensions*.^[6] For example, the matrix A above is a 3×2 matrix.

Matrices with a single row are called [row matrices](#) or *row vectors*, and those with a single column are called [column matrices](#) or *column vectors*. A matrix with the same number of rows and columns is called a [square matrix](#).^[7] A matrix with an infinite number of rows or columns (or both) is called an [infinite matrix](#). In some contexts, such as [computer algebra programs](#), it is useful to consider a matrix with no rows or no columns, called an [empty matrix](#).^[8]

Overview of a matrix size

Name	Size	Example	Description
Row matrix	$1 \times n$	$\begin{bmatrix} 3 & 7 & 2 \end{bmatrix}$	A matrix with one row and more than one columns, sometimes used to represent a vector
Column matrix	$n \times 1$	$\begin{bmatrix} 4 \\ 1 \\ 8 \end{bmatrix}$	A matrix with one column and more than one rows, sometimes used to represent a vector
Square matrix	$n \times n$	$\begin{bmatrix} 9 & 13 & 5 \\ 1 & 11 & 7 \\ 2 & 6 & 3 \end{bmatrix}$	A matrix with the same number of rows and columns, sometimes used to represent a linear transformation from a vector space to itself, such as reflection , rotation , or shearing .

Notation

The specifics of symbolic matrix notation vary widely, with some prevailing trends. Matrices are commonly written in [square brackets](#) or [parentheses](#),^[9] so that an $m \times n$ matrix A is represented as

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = \left(\begin{array}{cccc} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{array} \right).$$

This may be abbreviated by writing only a single generic term, possibly along with indices, as in

$$A = (a_{ij}), [a_{ij}], \text{ or } (a_{ij})_{1 \leq i \leq m, 1 \leq j \leq n} \text{ or } A = \mathbf{a}_{ij} \text{ in the case that } n = m.$$

Matrices are usually symbolized using [upper-case](#) letters (such as A in the examples above),^[10] while the corresponding [lower-case](#) letters, with two subscript indices (e.g., a_{11} , or $a_{1,1}$), represent the entries.^[11] In addition to using upper-case letters to symbolize matrices, many authors use a special [typographical style](#), commonly boldface roman (non-italic), to further distinguish matrices from other mathematical objects. An alternative notation involves the use of a double-underline with the variable name, with or without boldface style, as in $\underline{\underline{A}}$.^[12]

The entry in the i th row and j th column of a matrix $\mathbf{A}$ is sometimes referred to as the i, j or (i, j) entry of the matrix, and commonly denoted by $a_{i,j}$ or a_{ij} .^[13] Alternative notations for that entry are $A[i, j]$ and $A_{i,j}$. For example, the $(1,3)$ entry of the following matrix A is 5 (also denoted a_{13} , $a_{1,3}$, $A[1,3]$ or $A_{1,3}$):

$$A = \begin{bmatrix} 4 & -7 & 5 & 0 \\ -2 & 0 & 11 & 8 \\ 19 & 1 & -3 & 12 \end{bmatrix}$$

Sometimes, the entries of a matrix can be defined by a formula such as $a_{i,j} = f(i, j)$. For example, each of the entries of the following matrix A is determined by the formula $a_{ij} = i - j$.

$$A = \begin{bmatrix} 0 & -1 & -2 & -3 \\ 1 & 0 & -1 & -2 \\ 2 & 1 & 0 & -1 \end{bmatrix}$$

In this case, the matrix itself is sometimes defined by that formula, within square brackets or double parentheses. For example, the matrix above is defined as $A = [i - j]$ or $A = ((i - j))$. If matrix size is $m \times n$, the above-mentioned formula $f(i, j)$ is valid for any $i = 1, \dots, m$ and any $j = 1, \dots, n$. This can be specified separately or indicated using $m \times n$ as a subscript. For

instance, the matrix A above is 3×4 , and can be defined as $A=[i-j](i=1,2,3; j=1, \dots, 4)$ or $A=\begin{bmatrix} 0 & -1 & -2 & -3 \\ -1 & -2 & -3 & -4 \\ -2 & -3 & -4 & -5 \end{bmatrix}$.

Some programming languages utilize doubly subscripted arrays (or arrays of arrays) to represent an m -by- n matrix. Some programming languages start the numbering of array indexes at zero, in which case the entries of an $m \times n$ matrix are indexed by $0 \leq i \leq m-1$ and $0 \leq j \leq n-1$.^[14] This article follows the more common convention in mathematical writing where enumeration starts from 1.

The set of all m -by- n real matrices is often denoted $M(m, n)$, or $M_{m \times n}(R)$. The set of all $m \times n$ matrices over another field, or over a ring R , is similarly denoted $M(m, n, R)$, or $M_{m \times n}(R)$. If $m = n$, such as in the case of square matrices, one does not repeat the dimension: $M(n, R)$, or $M_n(R)$.^[15] Often, M , or Mat , is used in place of M .^[16]

Basic operations

Several basic operations can be applied to matrices. Some, such as transposition and submatrix do not depend on the nature of the entries. Others, such as matrix addition, scalar multiplication, matrix multiplication, and row operations involve operations on matrix entries and therefore require that matrix entries are numbers or belong to a field or a ring.^[17]

In this section, it is supposed that matrix entries belong to a fixed ring, which is typically a field of numbers.

Addition

Matrix addition and subtraction require matrices of a consistent size, and are calculated entrywise. The sum $\mathbf{A} + \mathbf{B}$ and the difference $\mathbf{A} - \mathbf{B}$ of two $m \times n$ matrices are:^[18]

$\begin{bmatrix} 1 & 3 & 1 \\ 1 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 5 \\ 7 & 5 & 0 \end{bmatrix} = \begin{bmatrix} 1+0 & 3+0 & 1+5 \\ 1+7 & 0+5 & 0+0 \end{bmatrix} = \begin{bmatrix} 1 & 3 & 6 \\ 8 & 5 & 0 \end{bmatrix}$

For example,

$$\begin{bmatrix} 1 & 3 & 1 \\ 1 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 5 \\ 7 & 5 & 0 \end{bmatrix} = \begin{bmatrix} 1+0 & 3+0 & 1+5 \\ 1+7 & 0+5 & 0+0 \end{bmatrix} = \begin{bmatrix} 1 & 3 & 6 \\ 8 & 5 & 0 \end{bmatrix}$$

Familiar properties of numbers extend to these operations on matrices: for example, addition is [commutative](#), that is, the matrix sum does not depend on the order of the summands: $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$.^[19]

Scalar multiplication

The product $c\mathbf{A}$ of a number c (also called a [scalar](#) in this context) and a matrix $\mathbf{A}$ is computed by multiplying each entry of $\mathbf{A}$ by c :^[20]

⚠ This operation is called *scalar multiplication*, but its result is not named "scalar product" to avoid confusion, since "scalar product" is often used as a synonym for "[inner product](#)".^[21] For example:

$$2 \cdot \begin{bmatrix} 1 & 8 & -3 \\ 4 & -2 & 5 \end{bmatrix} = \begin{bmatrix} 2 \cdot 1 & 2 \cdot 8 & 2 \cdot (-3) \\ 2 \cdot 4 & 2 \cdot (-2) & 2 \cdot 5 \end{bmatrix} = \begin{bmatrix} 2 & 16 & -6 \\ 8 & -4 & 10 \end{bmatrix}$$

Matrix subtraction is consistent with composition of matrix addition with scalar multiplication by -1 :^[22]

$$\mathbf{A} - \mathbf{B} = \mathbf{A} + (-1) \cdot \mathbf{B}$$

Transpose

The *transpose* of an $m \times n$ matrix $\mathbf{A}$ is the $n \times m$ matrix $\mathbf{A}^T$ (also denoted $\mathbf{A}^{\text{tr}}$ or ${}^t\mathbf{A}$) formed by turning rows into columns and vice versa:

$(\mathbf{A}^T)_{i,j} = \mathbf{A}_{j,i}$. For example:

$$\begin{bmatrix} 1 & 2 & 3 \\ 0 & -6 & 7 \end{bmatrix}^T = \begin{bmatrix} 1 & 0 \\ 2 & -6 \\ 3 & 7 \end{bmatrix}$$

The transpose is compatible with addition and scalar multiplication, as expressed by $(c\mathbf{A})^T = c(\mathbf{A}^T)$ and $(\mathbf{A} + \mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T$. Finally, $(\mathbf{A}^T)^T = \mathbf{A}$.^[23]

Matrix multiplication

Multiplication of two matrices corresponds to the composition of [linear transformations](#) represented by each matrix. It is defined if and only if the number of columns of the left matrix is the same as the number of rows of the right matrix. If $\mathbf{A}$ is an $m \times n$ matrix and $\mathbf{B}$ is an $n \times p$ matrix, then their *matrix product* $\mathbf{AB}$ is the $m \times p$ matrix whose entries are given by the [dot product](#) of the corresponding row of $\mathbf{A}$ and the corresponding column

of **B**:^[24]

where $1 \leq i \leq m$ and $1 \leq j \leq p$.^[25] For example, the underlined entry 2340 in the product is calculated as $(2 \times 1000) + (3 \times 100) + (4 \times 10) = 2340$:

$$\begin{bmatrix} 2 & 3 & 4 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & \underline{1000} \\ 1 & \underline{100} \\ 0 & \underline{10} \end{bmatrix} = \begin{bmatrix} 3 & \underline{2340} \\ 0 & 1000 \end{bmatrix}.$$

Matrix multiplication satisfies the rules $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$ ([associativity](#)), and $(\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC}$ as well as $\mathbf{C}(\mathbf{A} + \mathbf{B}) = \mathbf{CA} + \mathbf{CB}$ (left and right [distributivity](#)), whenever the size of the matrices is such that the various products are defined.^[26] The product $\mathbf{AB}$ may be defined without $\mathbf{BA}$ being defined, namely if $\mathbf{A}$ and $\mathbf{B}$ are $m \times n$ and $n \times k$ matrices, respectively, and $m \neq k$. Even if both products are defined, they generally need not be equal, that is:^[27]

$\mathbf{AB} \neq \mathbf{BA}$.

In other words, matrix multiplication is not [commutative](#), in marked contrast to (rational, real, or complex) numbers, whose product is independent of the order of the factors.^[24] An example of two matrices not commuting with each other is:

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 3 \end{bmatrix}, \text{ whereas } \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 3 & 4 \\ 0 & 0 \end{bmatrix}.$$

Besides the ordinary matrix multiplication just described, other less frequently used operations on matrices that can be considered forms of multiplication also exist, such as the [Hadamard product](#) and the [Kronecker product](#).^[28] They arise in solving matrix equations such as the [Sylvester equation](#).^[29]

Row operations

There are three types of row operations:^{[30][31]}

1. row addition, that is, adding a row to another.
2. row multiplication, that is, multiplying all entries of a row by a non-zero constant;
3. row switching, that is, interchanging two rows of a matrix;

These operations are used in several ways, including solving [linear equations](#) and finding [matrix inverses](#) with [Gauss elimination](#) and Gauss-Jordan elimination, respectively.^[32]

Submatrix

A **submatrix** of a matrix is a matrix obtained by deleting any collection of rows or columns or both. ^{[33][34][35]} For example, from the following 3×4 matrix, we can construct a 2×3 submatrix by removing row 3 and column 2:

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 3 & 4 \\ 5 & 7 & 8 \end{bmatrix}.$$

The [minors](#) and cofactors of a matrix are found by computing the [determinant](#) of certain submatrices. ^{[35][36]}

A **principal submatrix** is a square submatrix obtained by removing certain rows and columns. The definition varies from author to author. According to some authors, a principal submatrix is a submatrix in which the set of row indices that remain is the same as the set of column indices that remain. ^[37] ^[38] Other authors define a principal submatrix as one in which the first k rows and columns, for some number k , are the ones that remain; ^[39] this type of submatrix has also been called a **leading principal submatrix**. ^[40]

Linear equations

Matrices can be used to compactly write and work with multiple linear equations, that is, systems of linear equations. For example, if $\mathbf{A}$ is an $m \times n$ matrix, $\mathbf{x}$ designates a column vector (that is, $n \times 1$ matrix) of n variables $x_1, x_2, \dots, x_n$, and $\mathbf{b}$ is an $m \times 1$ column vector, then the matrix equation

$A\mathbf{x} = \mathbf{b}$ is equivalent to the system of linear equations ^[41]

$$\begin{aligned} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n &= b_1 \\ &\vdots \\ a_{m,1}x_1 + a_{m,2}x_2 + \cdots + a_{m,n}x_n &= b_m \end{aligned}$$

Using matrices, this can be solved more compactly than would be possible by writing out all the equations separately. If $n = m$ and the equations are [independent](#), then this can be done by writing ^[42]

$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$ where $\mathbf{A}^{-1}$ is the [inverse matrix](#) of $\mathbf{A}$. If $\mathbf{A}$ has no inverse, solutions—if any—can be found using its [generalized inverse](#). ^[43]

Linear transformations

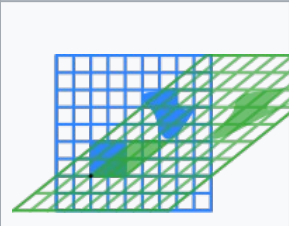
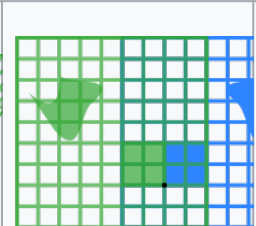
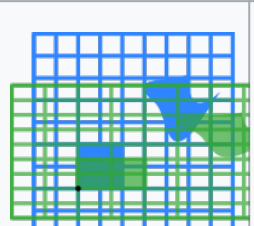
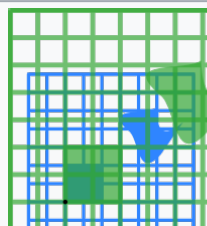
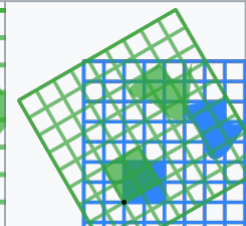
Matrices and matrix multiplication reveal their essential features when related to *linear transformations*, also known as *linear maps*. A real m -by- n matrix $\mathbf{A}$ gives rise to a linear transformation $R^n \rightarrow R^m$ mapping each vector $\mathbf{x}$ in R^n to the (matrix) product $\mathbf{Ax}$, which is a vector in R^m .

Conversely, each linear transformation $f: R^n \rightarrow R^m$ arises from a unique m -by- n matrix $\mathbf{A}$: explicitly, the (i, j) -entry of $\mathbf{A}$ is the i th coordinate of $f(\mathbf{e}_j)$, where $\mathbf{e}_j = (0, \dots, 0, 1, 0, \dots, 0)$ is the [unit vector](#) with 1 in the j th position and 0 elsewhere. The matrix $\mathbf{A}$ is said to represent the linear map f , and $\mathbf{A}$ is called the *transformation matrix* of f .^[44]

For example, the 2×2 matrix

$A = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$ can be viewed as the transform of the [unit square](#) into

a [parallelogram](#) with vertices at $(0, 0)$, (a, b) , $(a + c, b + d)$, and (c, d) . The parallelogram pictured at the right is obtained by multiplying $\mathbf{A}$ with each of the column vectors $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$, $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$, and $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ in turn. These vectors define the vertices of the unit square.^[45] The following table shows several 2×2 real matrices with the associated linear maps of R^2 . The [blue](#) original is mapped to the [green](#) grid and shapes. The origin $(0, 0)$ is marked with a black point.

Horizontal shear ^[46] with $m = 1.25$.	Reflection ^[47] through the vertical axis	Squeeze mapping ^[48] with $r = 3/2$	Scaling ^[49] by a factor of $3/2$	Rotation ^[48] by $\pi/6 = 30^\circ$
$\begin{bmatrix} 1 & 1.25 \\ 0 & 1 \end{bmatrix}$	$\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$	$\begin{bmatrix} \frac{3}{2} & 0 \\ 0 & \frac{2}{3} \end{bmatrix}$	$\begin{bmatrix} \frac{3}{2} & 0 \\ 0 & \frac{3}{2} \end{bmatrix}$	$\begin{bmatrix} \cos\left(\frac{\pi}{6}\right) & -\sin\left(\frac{\pi}{6}\right) \\ \sin\left(\frac{\pi}{6}\right) & \cos\left(\frac{\pi}{6}\right) \end{bmatrix}$
				

Under the [1-to-1 correspondence](#) between matrices and linear maps, matrix multiplication corresponds to [composition](#) of maps:^[50] if a k -by- m matrix $\mathbf{B}$ represents another linear map $g: R^m \rightarrow R^k$, then the composition $g \circ f$ is represented by $\mathbf{BA}$ since^[51]

$$(g \circ f)(x) = g(f(x)) = g(Ax) = B(Ax) = (BA)x.$$

The last equality follows from the above-mentioned associativity of matrix multiplication.

The [rank of a matrix](#) $\mathbf{A}$ is the maximum number of [linearly independent](#) row vectors of the matrix, which is the same as the maximum number of linearly independent column vectors.^[52] Equivalently it is the [dimension](#) of the [image](#) of the linear map represented by $\mathbf{A}$.^[53] The [rank-nullity theorem](#) states that the dimension of the [kernel](#) of a matrix plus the rank equals the number of columns of the matrix.^[54]

Square matrix

A [square matrix](#) is a matrix with the same number of rows and columns. An n -by- n matrix is known as a square matrix of order n . Any two square matrices of the same order can be added and multiplied. The entries a_{ii} form the [main diagonal](#) of a square matrix. They lie on the imaginary line running from the top left corner to the bottom right corner of the matrix.^[55]

Square matrices of a given dimension form a [noncommutative ring](#), which is one of the most common examples of a noncommutative ring.^[56]

Main types

Diagonal and triangular matrix

If all entries of $\mathbf{A}$ below the main diagonal are zero, $\mathbf{A}$ is called an *upper [triangular matrix](#)*. Similarly, if all entries of $\mathbf{A}$ above the main diagonal are zero, $\mathbf{A}$ is called a *lower triangular matrix*.^[57] If all entries outside the main diagonal are zero, $\mathbf{A}$ is called a [diagonal matrix](#).^[58]

Name	Example with $n = 3$
------	----------------------

Diagonal matrix	$\begin{bmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{bmatrix}$
Lower triangular matrix	$\begin{bmatrix} a_{11} & 0 & 0 \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$
Upper triangular matrix	$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{bmatrix}$

Identity matrix

The *identity matrix* $\mathbf{I}_n$ of size n is the n -by- n matrix in which all the elements on the [main diagonal](#) are equal to 1 and all other elements are equal to 0, ^[59] for example,

$$\begin{matrix} I_1 & \text{is} & [1], \\ I_2 & \text{is} & \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \\ \vdots & & \end{matrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \text{It is a square matrix of order } n, \text{ and also a}$$

special kind of [diagonal matrix](#). It is called an identity matrix because multiplication with it leaves a matrix unchanged: ^[59]

$\mathbf{A}\mathbf{I}_n = \mathbf{I}_m\mathbf{A} = \mathbf{A}$ for any m -by- n matrix $\mathbf{A}$.

A scalar multiple of an identity matrix is called a *scalar* matrix. ^[60]

Symmetric or skew-symmetric matrix

A square matrix $\mathbf{A}$ that is equal to its transpose, that is, $\mathbf{A} = \mathbf{A}^T$, is a [symmetric matrix](#). If instead, $\mathbf{A}$ is equal to the negative of its transpose, that is, $\mathbf{A} = -\mathbf{A}^T$, then $\mathbf{A}$ is a [skew-symmetric matrix](#). In complex matrices, symmetry is often replaced by the concept of [Hermitian matrices](#), which satisfies $\mathbf{A}^* = \mathbf{A}$, where the star or [asterisk](#) denotes the [conjugate transpose](#) of the matrix, that is, the transpose of the [complex conjugate](#) of $\mathbf{A}$. ^[61]

By the [spectral theorem](#), real symmetric matrices and complex Hermitian matrices have an [eigenbasis](#); that is, every vector is expressible as a [linear combination](#) of eigenvectors. In both cases, all eigenvalues are real. ^[62] This

theorem can be generalized to infinite-dimensional situations related to matrices with infinitely many rows and columns.^[63]

Invertible matrix and its inverse

A square matrix $\mathbf{A}$ is called [invertible](#) or *non-singular* if there exists a matrix $\mathbf{B}$ such that^{[64][65]}

$AB=BA=I_n$, where I_n is the $n \times n$ [identity matrix](#) with 1 for each entry on the [main diagonal](#) and 0 elsewhere. If $\mathbf{B}$ exists, it is unique and is called the [inverse matrix](#) of $\mathbf{A}$, denoted $\mathbf{A}^{-1}$.^[66]

There are many [algorithms](#) for testing whether a square matrix is invertible, and, if it is, computing its inverse. One of the oldest, which is still in common use is [Gaussian elimination](#).^[67]

Definite matrix

A symmetric real matrix $\mathbf{A}$ is called [positive-definite](#) if the associated [quadratic form](#)

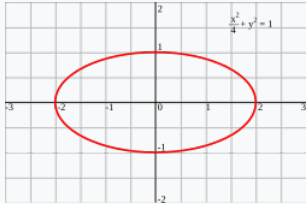
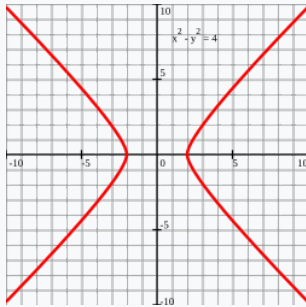
$f(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x}$ has a positive value for every nonzero vector $\mathbf{x}$ in R^n . If $f(\mathbf{x})$ yields only negative values then $\mathbf{A}$ is [negative-definite](#); if f does produce both negative and positive values then $\mathbf{A}$ is [indefinite](#).^[68] If the quadratic form f yields only non-negative values (positive or zero), the symmetric matrix is called *positive-semidefinite* (or if only non-positive values, then negative-semidefinite); hence the matrix is indefinite precisely when it is neither positive-semidefinite nor negative-semidefinite.^[69]

A symmetric matrix is positive-definite if and only if all its eigenvalues are positive, that is, the matrix is positive-semidefinite and it is invertible.^[70] The table at the right shows two possibilities for 2-by-2 matrices. The eigenvalues of a diagonal matrix are simply the entries along the diagonal,^[71] and so in these examples, the eigenvalues can be read directly from the matrices themselves. The first matrix has two eigenvalues that are both positive, while the second has one that is positive and another that is negative.

Allowing as input two different vectors instead yields the [bilinear form](#) associated to $\mathbf{A}$:^[72]

$$B_A(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T \mathbf{A} \mathbf{y}.$$

In the case of complex matrices, the same terminology and results apply, with *symmetric matrix*, *quadratic form*, *bilinear form*, and *transpose* $\mathbf{x}^T$ replaced respectively by [Hermitian matrix](#), [Hermitian form](#), [sesquilinear form](#), and [conjugate transpose](#) $\mathbf{x}^H$.^[73]

Positive definite matrix	Indefinite matrix
$\begin{bmatrix} \frac{1}{4} & 0 \\ 0 & 1 \end{bmatrix}$	$\begin{bmatrix} \frac{1}{4} & 0 \\ 0 & -\frac{1}{4} \end{bmatrix}$
$Q(x, y) = \frac{1}{4}x^2 + y^2$	$Q(x, y) = \frac{1}{4}x^2 - \frac{1}{4}y^2$
 Points such that $Q(x, y) = 1$ (Ellipse)	 Points such that $Q(x, y) = 1$ (Hyperbola)

Orthogonal matrix

An *orthogonal matrix* is a square matrix with [real](#) entries whose columns and rows are [orthogonal unit vectors](#) (that is, [orthonormal](#) vectors).

[74] Equivalently, a matrix **A** is orthogonal if its [transpose](#) is equal to its [inverse](#):

$A^T = A^{-1}$, which entails

$A^T A = A A^T = I_n$, where I_n is the [identity matrix](#) of size n . [75]

An orthogonal matrix **A** is necessarily [invertible](#) (with inverse $\mathbf{A}^{-1} = \mathbf{A}^T$), [unitary](#) ($\mathbf{A}^{-1} = \mathbf{A}^*$), and [normal](#) ($\mathbf{A}^* \mathbf{A} = \mathbf{A} \mathbf{A}^*$).

The [determinant](#) of any orthogonal matrix is either +1 or −1. A *special orthogonal matrix* is an orthogonal matrix with [determinant](#) +1. As a [linear transformation](#), every orthogonal matrix with determinant +1 is a pure [rotation](#) without reflection, i.e., the transformation preserves the orientation of the transformed structure, while every orthogonal matrix with determinant −1 reverses the orientation, i.e., is a composition of a pure [reflection](#) and a (possibly null) rotation. The identity matrices have determinant 1 and are pure rotations by an angle zero. [76]

The [complex](#) analog of an orthogonal matrix is a [unitary matrix](#).^[77]

Main operations

Trace

The [trace](#), $\text{tr}(\mathbf{A})$ of a square matrix $\mathbf{A}$ is the sum of its diagonal entries. While matrix multiplication is not commutative as mentioned [above](#), the trace of the product of two matrices is independent of the order of the factors:^[78]
 $\text{tr}(AB) = \text{tr}(BA)$. This is immediate from the definition of matrix multiplication:^[79]

$\text{tr}(AB) = \sum_{i=1}^m \sum_{j=1}^n a_{ij} b_{ji} = \text{tr}(BA)$. It follows that the trace of the product of more than two matrices is independent of [cyclic permutations](#) of the matrices; however, this does not in general apply for arbitrary permutations. For example, $\text{tr}(\mathbf{ABC}) \neq \text{tr}(\mathbf{BAC})$, in general.^[80] Also, the trace of a matrix is equal to that of its transpose,^[81] that is,
 $\text{tr}(A) = \text{tr}(A^T)$.

Determinant

The *determinant* of a square matrix $\mathbf{A}$ (denoted $\det(\mathbf{A})$ or $|\mathbf{A}|$) is a number encoding certain properties of the matrix. A matrix is invertible [if and only if](#) its determinant is nonzero.^[82] Its [absolute value](#) equals the area (in R^2) or volume (in R^3) of the image of the unit square (or cube), while its sign corresponds to the orientation of the corresponding linear map: the determinant is positive if and only if the orientation is preserved.^[83]

The determinant of 2×2 matrices is given by^[84]

$\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc$. The determinant of 3×3 matrices involves six terms ([rule of Sarrus](#)). The more lengthy [Leibniz formula](#) generalizes these two formulae to all dimensions.^[85]

The determinant of a product of square matrices equals the product of their determinants:

$\det(AB) = \det(A) \cdot \det(B)$, or using alternate notation:^[86]

$\det(AB) = \det(A) \cdot \det(B)$. Adding a multiple of any row to another row, or a multiple of any column to another column, does not change the determinant.

Interchanging two rows or two columns affects the determinant by multiplying it by -1 .^[87] Using these operations, any matrix can be transformed to a lower (or upper) triangular matrix, and for such matrices, the determinant equals the product of the entries on the main diagonal; this

provides a method to calculate the determinant of any matrix. Finally, the [Laplace expansion](#) expresses the determinant in terms of [minors](#), that is, determinants of smaller matrices.^[88] This expansion can be used for a recursive definition of determinants (taking as starting case the determinant of a 1×1 matrix, which is its unique entry, or even the determinant of a 0×0 matrix, which is 1), that can be seen to be equivalent to the Leibniz formula. Determinants can be used to solve [linear systems](#) using [Cramer's rule](#), where the division of the determinants of two related square matrices equates to the value of each of the system's variables.^[89]

Eigenvalues and eigenvectors

A number λ and a nonzero vector $\mathbf{v}$ satisfying $A\mathbf{v} = \lambda \mathbf{v}$ are called an *eigenvalue* and an *eigenvector* of $\mathbf{A}$, respectively.^[90]
^[91] The number λ is an eigenvalue of an $n \times n$ matrix $\mathbf{A}$ if and only if $(\mathbf{A} - \lambda \mathbf{I}_n)$ is not invertible, which is [equivalent](#) to^[92] $\det(\mathbf{A} - \lambda \mathbf{I}) = 0$. The polynomial $p_{\mathbf{A}}$ in an [indeterminate](#) X given by evaluation of the determinant $\det(X\mathbf{I}_n - \mathbf{A})$ is called the [characteristic polynomial](#) of $\mathbf{A}$. It is a [monic polynomial](#) of [degree](#) n . Therefore the polynomial equation $p_{\mathbf{A}}(\lambda) = 0$ has at most n different solutions, that is, eigenvalues of the matrix.^[93] They may be complex even if the entries of $\mathbf{A}$ are real.^[94] According to the [Cayley-Hamilton theorem](#), $p_{\mathbf{A}}(\mathbf{A}) = \mathbf{0}$, that is, the result of substituting the matrix itself into its characteristic polynomial yields the [zero matrix](#).^[95]

Computational aspects

Matrix calculations can be often performed with different techniques. Many problems can be solved by both direct algorithms and iterative approaches. For example, the eigenvectors of a square matrix can be obtained by finding a [sequence](#) of vectors $\mathbf{x}_n$ [converging](#) to an eigenvector when n tends to [infinity](#).^[96]

To choose the most appropriate algorithm for each specific problem, it is important to determine both the effectiveness and precision of all the available algorithms. The domain studying these matters is called [numerical linear algebra](#).^[97] As with other numerical situations, two main aspects are the [complexity](#) of algorithms and their [numerical stability](#).

Determining the complexity of an algorithm means finding [upper bounds](#) or estimates of how many elementary operations such as additions and multiplications of scalars are necessary to perform some algorithm, for

example, multiplication of matrices. Calculating the matrix product of two n -by- n matrices using the definition given above needs n^3 multiplications, since for any of the n^2 entries of the product, n multiplications are necessary. The [Strassen algorithm](#) outperforms this "naive" algorithm; it needs only $n^{2.807}$ multiplications.^[98] Theoretically faster but impractical [matrix multiplication algorithms](#) have been developed,^[99] as have speedups to this problem using [parallel algorithms](#) or [distributed computation](#) systems such as [MapReduce](#).^[100]

In many practical situations, additional information about the matrices involved is known. An important case concerns [sparse matrices](#), that is, matrices whose entries are mostly zero. There are specifically adapted algorithms for, say, solving linear systems $\mathbf{Ax} = \mathbf{b}$ for sparse matrices $\mathbf{A}$, such as the [conjugate gradient method](#).^[101]

An algorithm is, roughly speaking, numerically stable if little deviations in the input values do not lead to big deviations in the result. For example, one can calculate the inverse of a matrix by computing its [adjugate matrix](#):

$A^{-1} = \text{adj}(A) / \det(A)$. However, this may lead to significant rounding errors if the determinant of the matrix is very small. The [norm of a matrix](#) can be used to capture the [conditioning](#) of linear algebraic problems, such as computing a matrix's inverse.^[102]

Decomposition

There are several methods to render matrices into a more easily accessible form. They are generally referred to as *matrix decomposition* or *matrix factorization* techniques. These techniques are of interest because they can make computations easier.

The [LU decomposition](#) factors matrices as a product of lower ($\mathbf{L}$) and an upper [triangular matrices](#) ($\mathbf{U}$).^[103] Once this decomposition is calculated, linear systems can be solved more efficiently by a simple technique called [forward and back substitution](#). Likewise, inverses of triangular matrices are algorithmically easier to calculate. The *Gaussian elimination* is a similar algorithm; it transforms any matrix to [row echelon form](#).^[104] Both methods proceed by multiplying the matrix by suitable [elementary matrices](#), which correspond to [permuting rows or columns](#) and adding multiples of one row to another row. [Singular value decomposition](#) (SVD) expresses any matrix $\mathbf{A}$ as a product $\mathbf{UDV}^*$, where $\mathbf{U}$ and $\mathbf{V}$ are [unitary matrices](#) and $\mathbf{D}$ is a diagonal matrix.^[105]

The [eigendecomposition](#) or *diagonalization* expresses $\mathbf{A}$ as a product $\mathbf{VDV}^{-1}$, where $\mathbf{D}$ is a diagonal matrix and $\mathbf{V}$ is a suitable invertible matrix.^[106] If $\mathbf{A}$ can

be written in this form, it is called [diagonalizable](#). More generally, and applicable to all matrices, the Jordan decomposition transforms a matrix into [Jordan normal form](#), that is to say matrices whose only nonzero entries are the eigenvalues λ_1 to λ_n of $\mathbf{A}$, placed on the main diagonal and possibly entries equal to one directly above the main diagonal, as shown at the right. ^[107] Given the eigendecomposition, the n th power of $\mathbf{A}$ (that is, n -fold iterated matrix multiplication) can be calculated via $\mathbf{A}^n = \mathbf{V} \mathbf{\Lambda}^n \mathbf{V}^{-1}$ and the power of a diagonal matrix can be calculated by taking the corresponding powers of the diagonal entries, which is much easier than doing the exponentiation for $\mathbf{A}$ instead. This can be used to compute the [matrix exponential](#) $e^{\mathbf{A}}$, a need frequently arising in solving [linear differential equations](#), [matrix logarithms](#) and [square roots of matrices](#). ^[108] To avoid numerically [ill-conditioned](#) situations, further algorithms such as the [Schur decomposition](#) can be employed. ^[109]

Abstract algebraic aspects and generalizations

Matrices can be generalized in different ways. Abstract algebra uses matrices with entries in more general [fields](#) or even [rings](#), while linear algebra codifies properties of matrices in the notion of linear maps. It is possible to consider matrices with infinitely many columns and rows. Another extension is [tensors](#), which can be seen as higher-dimensional arrays of numbers, as opposed to vectors, which can often be realized as sequences of numbers, while matrices are rectangular or two-dimensional arrays of numbers. ^[110] Matrices, subject to certain requirements tend to form [groups](#) known as matrix groups. ^[111] Similarly under certain conditions matrices form [rings](#) known as [matrix rings](#). ^[112] Though the product of matrices is not in general commutative, certain matrices form [fields](#) sometimes called matrix fields. ^[113] (However the term "matrix field" is ambiguous, also referring to certain forms of physical [fields](#) that continuously map points of some space to matrices. ^[114]) In general, matrices over any ring and their [multiplication](#) can be represented as the arrows and composition of arrows in a [category](#), the [category of matrices](#) over that ring. The objects of this category are natural numbers, representing the dimensions of the matrices. ^[115]

Matrices with entries in a field or ring

This article focuses on matrices whose entries are real or complex numbers. However, matrices can be considered with much more general types of entries than real or complex numbers. As a first step of generalization, any [field](#), that is, a [set](#) where [addition](#), [subtraction](#), [multiplication](#), and [division](#) operations are defined and well-behaved, may be used instead of R or C , for example [rational numbers](#) or [finite fields](#). For example, [coding theory](#) makes use of matrices over finite fields.^[116] Wherever [eigenvalues](#) are considered, as these are roots of a polynomial, they may exist only in a larger field than that of the entries of the matrix. For instance, they may be complex in the case of a matrix with real entries. The possibility to reinterpret the entries of a matrix as elements of a larger field (for example, to view a real matrix as a complex matrix whose entries happen to be all real) then allows considering each square matrix to possess a full set of eigenvalues.^[117] Alternatively one can consider only matrices with entries in an [algebraically closed field](#), such as C , from the outset.^[118]

Matrices whose entries are [polynomials](#),^[119] and more generally, matrices with entries in a [ring](#) R are widely used in mathematics.^[1] Rings are a more general notion than fields in that a division operation need not exist. The very same addition and multiplication operations of matrices extend to this setting, too. The set $M(n, R)$ (also denoted $M_n(R)$ ^[15]) of all square n -by- n matrices over R is a ring called [matrix ring](#), isomorphic to the [endomorphism ring](#) of the left R -[module](#) R^n .^[120] If the ring R is [commutative](#), that is, its multiplication is commutative, then the ring $M(n, R)$ is also an [associative algebra](#) over R . The [determinant](#) of square matrices over a commutative ring R can still be defined using the [Leibniz formula](#); such a matrix is invertible if and only if its determinant is [invertible](#) in R , generalizing the situation over a field F , where every nonzero element is invertible.^[121] Matrices over [superrings](#) are called [supermatrices](#).^[122]

Matrices do not always have all their entries in the same ring – or even in any ring at all. One special but common case is [block matrices](#), which may be considered as matrices whose entries themselves are matrices. The entries need not be square matrices, and thus need not be members of any [ring](#); but in order to multiply them, their sizes must fulfill certain conditions: each pair of submatrices that are multiplied in forming the overall product must have compatible sizes.^[123]

Relationship to linear maps

Linear maps $R^n \rightarrow R^m$ are equivalent to m -by- n matrices, as described [above](#). More generally, any linear map $f: V \rightarrow W$ between finite-[dimensional vector](#)

[spaces](#) can be described by a matrix $\mathbf{A} = (a_{ij})$, after choosing [bases](#) $\mathbf{v}_1, \dots, \mathbf{v}_n$ of V , and $\mathbf{w}_1, \dots, \mathbf{w}_m$ of W (so n is the dimension of V and m is the dimension of W), which is such that

$f(\mathbf{v}_j) = \sum_{i=1}^m a_{i,j} \mathbf{w}_i$ for $j=1, \dots, n$. In other words, column j of $\mathbf{A}$ expresses the

image of $\mathbf{v}_j$ in terms of the basis vectors $\mathbf{w}_i$ of W ; thus this relation uniquely determines the entries of the matrix $\mathbf{A}$. The matrix depends on the choice of the bases: different choices of bases give rise to different, but [equivalent matrices](#).^[124] Many of the above concrete notions can be reinterpreted in this light, for example, the transpose matrix $\mathbf{A}^T$ describes the [transpose of the linear map](#) given by $\mathbf{A}$, concerning the [dual bases](#).^[125]

These properties can be restated more naturally: the [category of matrices](#) with entries in a field k with multiplication as composition is [equivalent](#) to the category of finite-dimensional [vector spaces](#) and linear maps over this field.^[126]

More generally, the set of $m \times n$ matrices can be used to represent the R -linear maps between the free modules R^m and R^n for an arbitrary ring R with unity. When $n = m$ composition of these maps is possible, and this gives rise to the [matrix ring](#) of $n \times n$ matrices representing the [endomorphism ring](#) of R^n .^[127]

Matrix groups

A [group](#) is a mathematical structure consisting of a set of objects together with a [binary operation](#), that is, an operation combining any two objects to a third, subject to certain requirements.^[128] A group in which the objects are [invertible](#) $n \times n$ matrices and the group operation is matrix multiplication is called a *matrix group* of degree n .^[129] Every such matrix group is a [subgroup](#) of (that is, a smaller group contained within) the group of *all* invertible $n \times n$ matrices, the [general linear group](#) of degree n .^[130]

Any property of square matrices that is preserved under matrix products and inverses can be used to define a matrix group. For example, the set of all $n \times n$ matrices whose determinant is 1 form a group called the [special linear group](#) of degree n .^[131] The set of [orthogonal matrices](#), determined by the condition

$M^T M = I$, form the [orthogonal group](#).^[132] Every orthogonal matrix has [determinant](#) 1 or -1 . Orthogonal matrices with determinant 1 form a group called the *special orthogonal group*.^[133]

Every [finite group](#) is [isomorphic](#) to a matrix group, as one can see by considering the [regular representation](#) of the [symmetric group](#).^[134] General

groups can be studied using matrix groups, which are comparatively well understood, using [representation theory](#).^[135]

Infinite matrices

It is also possible to consider matrices with infinitely many rows and columns.^[136] The basic operations introduced above are defined the same way in this case. Matrix multiplication, however, and all operations stemming therefrom are only meaningful when restricted to certain matrices, since the [sum featuring in the above definition](#) of the matrix product will contain an infinity of summands.^[137] An easy way to circumvent this issue is to restrict to *finitary matrices* all of whose rows (or columns) contain only finitely many nonzero terms.^[138] As in the finite case (see [above](#)), where matrices describe linear maps, infinite matrices can be used to describe [operators on Hilbert spaces](#), where convergence and [continuity](#) questions arise. However, the explicit point of view of matrices tends to obfuscate the matter,^[139] and the abstract and more powerful tools of [functional analysis](#) are used instead, by relating matrices to linear maps (as in the finite case [above](#)), but imposing additional convergence and continuity constraints.

Empty matrix

An *empty matrix* is a matrix in which the number of rows or columns (or both) is zero.^{[140][8]} Empty matrices can be a useful [base case](#) for certain [recursive](#) constructions,^[141] and can help to deal with maps involving the [zero vector space](#).^[142] For example, if **A** is a 3×0 matrix and **B** is a 0×3 matrix, then **AB** is the 3×3 [zero matrix](#) corresponding to the null map from a 3-dimensional space V to itself, while **BA** is a 0×0 matrix. There is no common notation for empty matrices, but most [computer algebra systems](#) allow creating and computing with them.^[143] The determinant of the 0×0 matrix is conventionally defined to be 1, consistent with the [empty product](#) occurring in the Leibniz formula for the determinant.^[144] This value is also needed for consistency with the 2×2 case of the [Desnanot-Jacobi identity](#) relating determinants to the determinants of smaller matrices.^[145]

Matrices with entries in a semiring

A [semiring](#) is similar to a ring, but elements need not have [additive inverses](#), therefore one cannot do subtraction freely there. The definition of addition and multiplication of matrices with entries in a ring applies to matrices with entries in a semiring without modification. Matrices of fixed size with entries in a semiring form a [commutative monoid](#) $Mat(m, n; R)$ under addition.

^[146] Square matrices of fixed size with entries in a semiring form a semiring $Mat(n; R)$ under addition and multiplication. ^[146]

The determinant of an $n \times n$ square matrix M with entries in a [commutative semiring](#) R cannot be defined in general because the definition would involve additive inverses of semiring elements. What plays its role instead is the pair of positive and negative determinants

$$det_{+} M = \sum_{\sigma \in Alt(n)} M_{1\sigma(1)} \cdots M_{n\sigma(n)}$$

$$det_{-} M = \sum_{\sigma \in Sym(n) \setminus Alt(n)} M_{1\sigma(1)} \cdots M_{n\sigma(n)}$$

where the sums are taken over [even permutations](#) and odd permutations, respectively. ^{[147][148]}

Matrices with entries in a category

Matrices and their multiplication can be defined with entries objects of a [category](#) equipped with a "[tensor product](#)" similar to multiplication in a ring, having [coproducts](#) similar to addition in a ring, in that the former is [distributive](#) over the latter. ^[149] However, the multiplication thus defined may be only associative in a sense weaker than usual. These are part of a bigger structure called the *bicategory of matrices*. The complete description of the above summary for interested readers follows.

Let $(C, \otimes, I)$ be a [monoidal category](#) satisfying the following two conditions:

- All (small) [coproducts](#) exist; in particular, let $\emptyset$ be an [initial object](#).
- The functor $\otimes$ is distributive over coproducts; i.e., for all object X and a family of objects i in C , the canonical C -morphisms

$$\coprod_{i \in I} (X \otimes i Y_i) \rightarrow X \otimes \coprod_{i \in I} Y_i \quad \coprod_{i \in I} (i Y_i \otimes X) \rightarrow \left(\coprod_{i \in I} Y_i \right) \otimes X$$

are [isomorphisms](#). In particular, the canonical morphisms $\emptyset \rightarrow X \otimes \emptyset$ and $\emptyset \rightarrow \emptyset \otimes X$ are isomorphisms.

Then, the [bicategory](#) of C -matrices $Mat(C)$ is as follows: ^[149]

- The objects are the sets.

- A [1-morphism](#) $M: A \rightarrow B$ is a map $M: A \times B \rightarrow \text{Ob}(C)$; this is just a matrix over C .
 - The composition of 1-morphisms $M: A \rightarrow B$ and $N: B \rightarrow C$, which can be understood as matrix multiplication, is

$$(N \circ M)(a, c) = \coprod_{b \in B} M(a, b) \otimes N(b, c).$$
 - The identity 1-morphism on A is

$$id_A(a, b) = \begin{cases} I & a = b \\ \emptyset & a \neq b \end{cases}.$$
- A 2-morphism between 1-morphisms $M, N: A \rightarrow B$ is a family of C -morphisms i . The definition of vertical and horizontal composition of 2-morphisms is natural: the vertical composition is componentwise composition of C -morphisms; the horizontal composition is that derived from the functoriality of $\otimes$ and the [universal property](#) of coproducts.

In general, the bicategory of matrices need not be a strict [2-category](#). For example, the composition of 1-morphisms may not be associative in the usual strict sense, but only up to [coherent](#) isomorphism.

Applications

There are numerous applications of matrices, both in mathematics and other sciences. Some of them merely take advantage of the compact representation of a set of numbers in a matrix. For example, [Text mining](#) and automated [thesaurus](#) compilation makes use of [document-term matrices](#) such as [tf-idf](#) to track frequencies of certain words in several documents.^[150]

Complex numbers can be represented by particular real 2-by-2 matrices via $a + ib \leftrightarrow \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$, under which addition and multiplication of complex numbers and matrices correspond to each other. For example, 2-by-2 rotation matrices represent the multiplication with some complex number of [absolute value](#) 1, as [above](#). A similar interpretation is possible for [quaternions](#)^[151] and [Clifford algebras](#) in general.^[152]

In [game theory](#) and [economics](#), the [payoff matrix](#) encodes the payoff for two players, depending on which out of a given (finite) set of strategies the players choose.^[153] The expected outcome of the game, when both players play [mixed strategies](#), is obtained by multiplying this matrix on both sides by vectors representing the strategies.^[154] The [minimax theorem](#) central to

game theory is closely related to the [duality theory of linear programs](#), which are often formulated in terms of matrix-vector products.^[155]

Early [encryption](#) techniques such as the [Hill cipher](#) also used matrices. However, due to the linear nature of matrices, these codes are comparatively easy to break.^[156] [Computer graphics](#) uses matrices to represent objects; to calculate transformations of objects using affine [rotation matrices](#) to accomplish tasks such as projecting a three-dimensional object onto a two-dimensional screen, corresponding to a theoretical camera observation; and to apply image convolutions such as sharpening, blurring, edge detection, and more.^[157] Matrices over a [polynomial ring](#) are important in the study of [control theory](#).^[158]

[Chemistry](#) makes use of matrices in various ways, particularly since the use of [quantum theory](#) to discuss [molecular bonding](#) and [spectroscopy](#). Examples are the [overlap matrix](#) and the [Fock matrix](#) used in solving the [Roothaan equations](#) to obtain the [molecular orbitals](#) of the [Hartree–Fock method](#).^[159]

Graph theory

The [adjacency matrix](#) of a [finite graph](#) is a basic notion of [graph theory](#).^[160] It records which vertices of the graph are connected by an edge. Matrices containing just two different values (1 and 0 meaning for example "yes" and "no", respectively) are called [logical matrices](#). The [distance \(or cost\) matrix](#) contains information about the distances of the edges.^[161] These concepts can be applied to [websites](#) connected by [hyperlinks](#),^[162] or cities connected by roads etc., in which case (unless the connection network is extremely dense) the matrices tend to be [sparse](#), that is, contain few nonzero entries. Therefore, specifically tailored matrix algorithms can be used in [network theory](#).^[163]

Analysis and geometry

The [Hessian matrix](#) of a [differentiable function](#) $f: R^n \rightarrow R$ consists of the [second derivatives](#) of f concerning the several coordinate directions, that is,^[164]

$$H(f) = \left[\frac{\partial^2 f}{\partial x_i \partial x_j} \right].$$

It encodes information about the local growth behavior of the function: given a [critical point](#) $\mathbf{x} = (x_1, \dots, x_n)$, that is, a point where the first [partial](#)

[derivatives](#) $\partial f / \partial x_i$ of f vanish, the function has a [local minimum](#) if the Hessian matrix is [positive definite](#). [Quadratic programming](#) can be used to find global minima or maxima of quadratic functions closely related to the ones attached to matrices (see [above](#)).^[165]

Another matrix frequently used in geometrical situations is the [Jacobi matrix](#) of a differentiable map $f: R^n \rightarrow R^m$. If $f_1, \dots, f_m$ denote the components of f , then the Jacobi matrix is defined as^[166]

$$J(f) = \left[\frac{\partial f_i}{\partial x_j} \right]_{1 \leq i \leq m, 1 \leq j \leq n}$$
. If $n > m$, and if the rank of the Jacobi matrix attains its maximal value m , f is locally invertible at that point, by the [implicit function theorem](#).^[167]

[Partial differential equations](#) can be classified by considering the matrix of coefficients of the highest-order differential operators of the equation. For [elliptic partial differential equations](#) this matrix is positive definite, which has a decisive influence on the set of possible solutions of the equation in question.^[168]

The [finite element method](#) is an important numerical method to solve partial differential equations, widely applied in simulating complex physical systems. It attempts to approximate the solution to some equation by piecewise linear functions, where the pieces are chosen concerning a sufficiently fine grid, which in turn can be recast as a matrix equation.^[169]

Probability theory and statistics

[Stochastic matrices](#) are square matrices whose rows are [probability vectors](#), that is, whose entries are non-negative and sum up to one. Stochastic matrices are used to define [Markov chains](#) with finitely many states.^[170] A row of the stochastic matrix gives the probability distribution for the next position of some particle currently in the state that corresponds to the row. Properties of the Markov chain—like [absorbing states](#), that is, states that any particle attains eventually—can be read off the eigenvectors of the transition matrices.^[171]

Statistics also makes use of matrices in many different forms.^[172] [Descriptive statistics](#) is concerned with describing data sets, which can often be represented as [data matrices](#), which may then be subjected to [dimensionality reduction](#) techniques. The [covariance matrix](#) encodes the mutual [variance](#) of several [random variables](#).^[173] Another technique using matrices are [linear least squares](#), a method that approximates a finite set of pairs $(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)$, by a linear function

$y_i \approx a x_i + b, i=1, \dots, N$ which can be formulated in terms of matrices, related to the [singular value decomposition](#) of matrices.^[174]

[Random matrices](#) are matrices whose entries are random numbers, subject to suitable [probability distributions](#), such as [matrix normal distribution](#). Beyond probability theory, they are applied in domains ranging from [number theory](#) to [physics](#).^{[175][176]}

Quantum mechanics and particle physics

The first model of [quantum mechanics](#) ([Heisenberg](#), 1925) used infinite-dimensional matrices to define the operators that took over the role of variables like position, momentum and energy from classical physics.^[177] (This is sometimes referred to as [matrix mechanics](#).^[178]) Matrices, both finite and infinite-dimensional, have since been employed for many purposes in quantum mechanics. One particular example is the [density matrix](#), a tool used in calculating the [probabilities](#) of the outcomes of [measurements](#) performed on [physical systems](#).^{[179][180]}

Linear transformations and the associated [symmetries](#) play a [key role in modern physics](#). For example, [elementary particles](#) in [quantum field theory](#) are classified as representations of the [Lorentz group](#) of special relativity and, more specifically, by their behavior under the [spin group](#). Concrete representations involving the [Pauli matrices](#) and more general [gamma matrices](#) are an integral part of the physical description of [fermions](#), which behave as [spinors](#).^[181] For the three lightest [quarks](#), there is a group-theoretical representation involving the [special unitary group](#) $SU(3)$; for their calculations, physicists use a convenient matrix representation known as the [Gell-Mann matrices](#), which are also used for the $SU(3)$ [gauge group](#) that forms the basis of the modern description of strong nuclear interactions, [quantum chromodynamics](#). The [Cabibbo-Kobayashi-Maskawa matrix](#), in turn, expresses the fact that the basic quark states that are important for [weak interactions](#) are not the same as, but linearly related to the basic quark states that define particles with specific and distinct [masses](#).^[182]

Another matrix serves as a key tool for describing the scattering experiments that form the cornerstone of experimental particle physics: Collision reactions such as occur in [particle accelerators](#), where non-interacting particles head towards each other and collide in a small interaction zone, with a new set of non-interacting particles as the result, can be described as the scalar product of outgoing particle states and a linear combination of ingoing particle states. The linear combination is given by a matrix known as

the [S-matrix](#), which encodes all information about the possible interactions between particles.^[183]

Normal modes

A general application of matrices in physics is the description of linearly coupled harmonic systems. The [equations of motion](#) of such systems can be described in matrix form, with a mass matrix multiplying a generalized velocity to give the kinetic term, and a [force](#) matrix multiplying a displacement vector to characterize the interactions. The best way to obtain solutions is to determine the system's [eigenvectors](#), its [normal modes](#), by diagonalizing the matrix equation. Techniques like this are crucial when it comes to the internal dynamics of [molecules](#): the internal vibrations of systems consisting of mutually bound component atoms.^[184] They are also needed for describing mechanical vibrations, and oscillations in electrical circuits.^[185]

Geometrical optics

[Geometrical optics](#) provides further matrix applications. In this approximative theory, the [wave nature](#) of light is neglected. The result is a model in which [light rays](#) are indeed [geometrical rays](#). If the deflection of light rays by optical elements is small, the action of a [lens](#) or reflective element on a given light ray can be expressed as multiplication of a two-component vector with a two-by-two matrix called [ray transfer matrix analysis](#): the vector's components are the light ray's slope and its distance from the optical axis, while the matrix encodes the properties of the optical element. There are two kinds of matrices, viz. a *refraction matrix* describing the refraction at a lens surface, and a *translation matrix*, describing the translation of the plane of reference to the next refracting surface, where another refraction matrix applies. The optical system, consisting of a combination of lenses and reflective elements, is simply described by the matrix resulting from the product of the components' matrices.^[186]

The [Jones calculus](#) models the [polarization](#) of a light source as a 2×2 vector, and the effects of [optical filters](#) on this polarization vector as a matrix.^[48]

Electronics

Electronic circuits that are composed of linear components (such as resistors, inductors and capacitors) obey [Kirchhoff's circuit laws](#), which leads to a system of linear equations, which can be described with a matrix equation that relates the source currents and voltages to the resultant currents and voltages at each point in the circuit, and where the matrix entries are determined by the circuit.^[187]

History

Matrices have a long history of application in solving [linear equations](#) but they were known as arrays until the 1800s. The [Chinese text *The Nine Chapters on the Mathematical Art*](#) written in the 10th–2nd century BCE is the first example of the use of array methods to solve [simultaneous equations](#),^[188] including the concept of [determinants](#). In 1545 Italian mathematician [Gerolamo Cardano](#) introduced the method to Europe when he published *Ars Magna*.^[189] The [Japanese mathematician Seki](#) used the same array methods to solve simultaneous equations in 1683.^[190] The Dutch mathematician [Jan de Witt](#) represented transformations using arrays in his 1659 book *Elements of Curves* (1659).^[191] Between 1700 and 1710 [Gottfried Wilhelm Leibniz](#) publicized the use of arrays for recording information or solutions and experimented with over 50 different systems of arrays.^[189] [Cramer](#) presented [his rule](#) in 1750.^{[192][193]}

This use of the term *matrix* in mathematics (an English word for "womb" in the 19th century, from Latin, as well as a jargon word [in printing](#), [in biology](#) and [in geology](#)^[194]) was coined by [James Joseph Sylvester](#) in 1850,^[195] who understood a matrix as an object giving rise to several determinants today called [minors](#), that is to say, determinants of smaller matrices that derive from the original one by removing columns and rows. In an 1851 paper, Sylvester explains:^[196]

I have in previous papers defined a "Matrix" as a rectangular array of terms, out of which different systems of determinants may be engendered from the womb of a common parent.

[Arthur Cayley](#) published a treatise on geometric transformations using matrices that were not rotated versions of the coefficients being investigated as had previously been done. Instead, he defined operations such as addition, subtraction, multiplication, and division as transformations of those matrices and showed the associative and distributive properties held. Cayley investigated and demonstrated the non-commutative property of matrix multiplication as well as the commutative property of matrix addition.^[189] Early matrix theory had limited the use of arrays almost exclusively to

determinants and Cayley's abstract matrix operations were revolutionary. He was instrumental in proposing a matrix concept independent of equation systems. In 1858, Cayley published his *A memoir on the theory of matrices*^{[197][198]} in which he proposed and demonstrated the [Cayley-Hamilton theorem](#).^[189]

The English mathematician [Cuthbert Edmund Cullis](#) was the first to use modern bracket notation for matrices in 1913 and he simultaneously demonstrated the first significant use of the notation $\mathbf{A} = [a_{i,j}]$ to represent a matrix where $a_{i,j}$ refers to the i th row and the j th column.^[189]

The modern study of determinants sprang from several sources.^[199] [Number-theoretical](#) problems led [Gauss](#) to relate coefficients of [quadratic forms](#), that is, expressions such as $x^2 + xy - 2y^2$, and [linear maps](#) in three dimensions to matrices. [Eisenstein](#) further developed these notions, including the remark that, in modern parlance, [matrix products](#) are [non-commutative](#). [Cauchy](#) was the first to prove general statements about determinants, using as the definition of the determinant of a matrix $\mathbf{A} = [a_{i,j}]$ the following: replace the powers a_j^k by $a_{j,k}$ in the [polynomial](#)

$a_1 a_2 \cdots a_n \prod_{i < j} (a_j - a_i)$, where $\prod$ denotes the [product](#) of the indicated terms. He also showed, in 1829, that the [eigenvalues](#) of symmetric matrices are real.^[200] [Jacobi](#) studied "functional determinants"—later called [Jacobi determinants](#) by Sylvester—which can be used to describe geometric transformations at a local (or [infinitesimal](#)) level, see [above](#). [Kronecker](#)'s *Vorlesungen über die Theorie der Determinanten*^[201] and [Weierstrass](#)'s *Zur Determinantentheorie*,^[202] both published in 1903, first treated determinants [axiomatically](#), as opposed to previous more concrete approaches such as the mentioned formula of Cauchy. At that point, determinants were firmly established.^{[203][199]}

Many theorems were first established for small matrices only, for example, the [Cayley-Hamilton theorem](#) was proved for 2×2 matrices by Cayley in the aforementioned memoir, and by [Hamilton](#) for 4×4 matrices. [Frobenius](#), working on [bilinear forms](#), generalized the theorem to all dimensions (1898). Also at the end of the 19th century, the [Gauss-Jordan elimination](#) (generalizing a special case now known as [Gauss elimination](#)) was established by [Wilhelm Jordan](#). In the early 20th century, matrices attained a central role in linear algebra,^[204] partially due to their use in the classification of the [hypercomplex number](#) systems of the previous century.^[205]

The inception of [matrix mechanics](#) by [Heisenberg](#), [Born](#) and [Jordan](#) led to studying matrices with infinitely many rows and columns.^[206] Later, [von Neumann](#) carried out the [mathematical formulation of quantum mechanics](#), by further developing [functional analytic](#) notions such as [linear](#)

[operators](#) on [Hilbert spaces](#), which, very roughly speaking, correspond to [Euclidean space](#), but with an infinity of [independent directions](#).^[207]

Other historical usages of the word "matrix" in mathematics

The word has been used in unusual ways by at least two authors of historical importance.

[Bertrand Russell](#) and [Alfred North Whitehead](#) in their *[Principia Mathematica](#)* (1910–1913) use the word "matrix" in the context of their [axiom of reducibility](#). They proposed this axiom as a means to reduce any function to one of lower type, successively, so that at the "bottom" (0 order) the function is identical to its [extension](#).^[208]

Let us give the name of *matrix* to any function, of however many variables, that does not involve any [apparent variables](#). Then, any possible function other than a matrix derives from a matrix using generalization, that is, by considering the proposition that the function in question is true with all possible values or with some value of one of the arguments, the other argument or arguments remaining undetermined.

For example, a function $\Phi(x, y)$ of two variables x and y can be reduced to a *collection* of functions of a single variable, such as y , by "considering" the function for all possible values of "individuals" a_i substituted in place of a variable x . And then the resulting collection of functions of the single variable y , that is, $\forall a_i: \Phi(a_i, y)$, can be reduced to a "matrix" of values by "considering" the function for all possible values of "individuals" b_j substituted in place of variable y :
 $\forall b_j \forall a_i: \phi(a_i, b_j)$.

[Alfred Tarski](#) in his 1941 *Introduction to Logic* used the word "matrix" synonymously with the notion of [truth table](#) as used in mathematical logic.^[209]

See also

- [List of named matrices](#)
- [Gram-Schmidt process](#) – Orthonormalization of a set of vectors
- [Irregular matrix](#)

- [Matrix calculus](#) – Specialized notation for multivariable calculus
- [Matrix function](#) – Function that maps matrices to matrices

Read more (what I didn't insert back then)

Matrix (mathematics)

$$\begin{matrix}
 & \begin{matrix} 1 & 2 & \cdots & n \end{matrix} \\
 \begin{matrix} 1 \\ 2 \\ 3 \\ \vdots \\ m \end{matrix} & \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ a_{31} & a_{32} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}
 \end{matrix}$$

An $m \times n$ matrix: the m rows are horizontal and the n columns are vertical. Each element of a matrix is often denoted by a variable with two [subscripts](#). For example, $a_{2,1}$ represents the element at the second row and first column of the matrix.

Basic operations

Addition

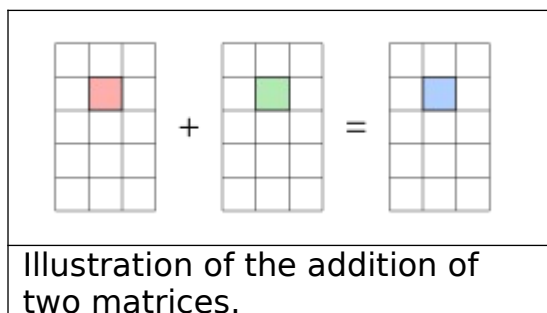
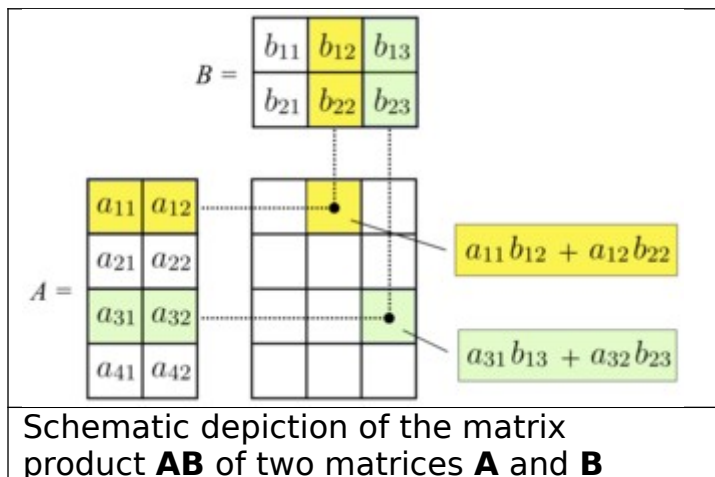
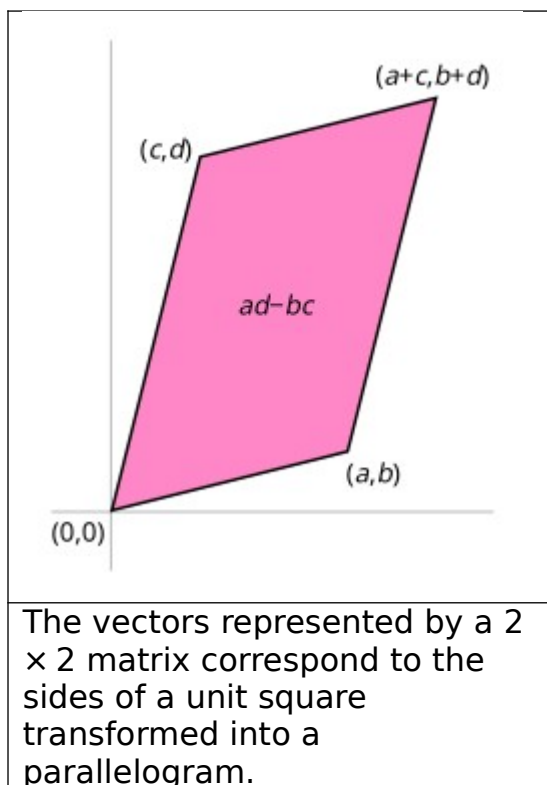


Illustration of the addition of two matrices.

Matrix multiplication



Linear transformations



Square matrix

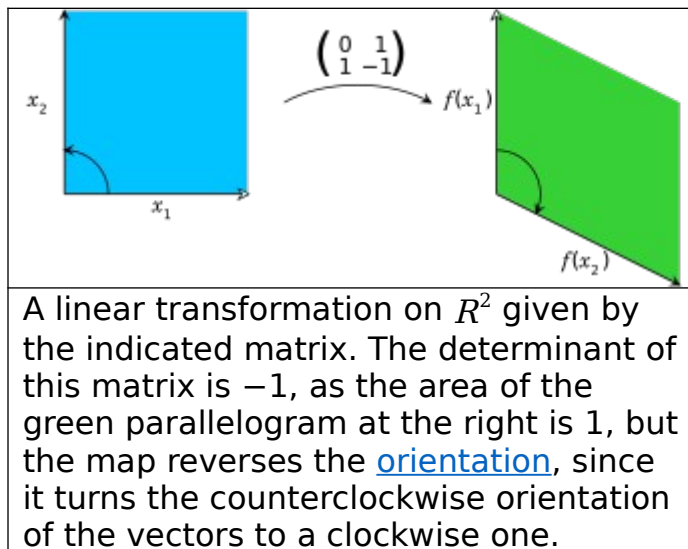
Main types

Diagonal and triangular matrix

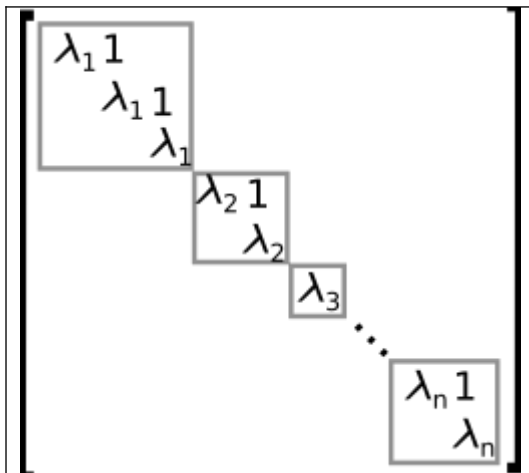
Name	Example with $n = 3$
Diagonal matrix	$\begin{bmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{bmatrix}$
Lower triangular matrix	$\begin{bmatrix} a_{11} & 0 & 0 \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$
Upper triangular matrix	$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{bmatrix}$

Main operations

Determinant



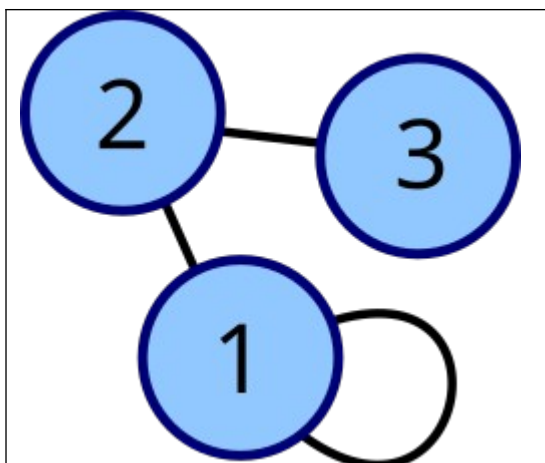
Decomposition



An example of a matrix in Jordan normal form. The grey blocks are called Jordan blocks.

Applications

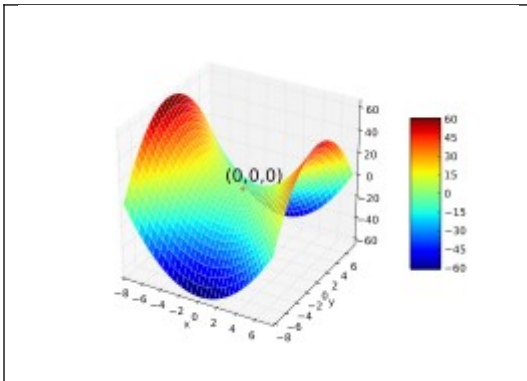
Graph theory



An undirected graph with adjacency matrix:

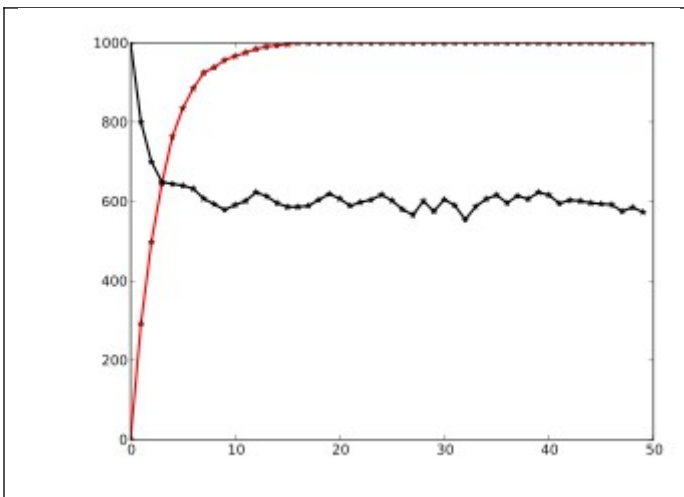
$$\begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}.$$

Analysis and geometry



At the [saddle point](#) ($x = 0, y = 0$) (red) of the function $f(x, y) = x^2 - y^2$, the Hessian matrix $\begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix}$ is [indefinite](#).

Probability theory and statistics



Two different Markov chains. The chart depicts the number of particles (of a total of 1000) in state "2". Both limiting values can be determined from the transition matrices, which are given by $\begin{bmatrix} 0.7 & 0 \\ 0.3 & 1 \end{bmatrix}$ (red) and $\begin{bmatrix} 0.7 & 0.2 \\ 0.3 & 0.8 \end{bmatrix}$ (black).